Aviation Research Lab
Institute of Aviation

University of Illinois
at Urbana-Champaign
1 Airport Road
Savoy, Illinois 61874

PILOT EXPECTANCY AND ATTENTIONAL EFFECTS FOR HAZARD AWARENESS

Christopher D. Wickens

Final Technical Report ARL-00-6/FAA-00-3

July 2000

Prepared for

Federal Aviation Administration
Civil Aeromedical Institute
Oklahoma City, OK

Contract DTFA 98-G-022

PILOT EXPECTANCY AND ATTENTIONAL EFFECTS FOR HAZARD AWARENESS

FINAL REPORT: FAA

EXECUTIVE SUMMARY

This research examines the implications of three projected technological innovations on attention and navigational hazard awareness for the general aviation pilot, comparing these against baseline performance. In Experiments 1 and 2 we examine the joint implications of freeflight, availed by satellite navigation technology and the Cockpit Display of Traffic Information (CDTI), for the pilots' decision and attentional strategies, employed to avoid traffic hazards. In Experiment 3 we address the implications of digital datalink on pilot traffic awareness and communications reliability, examining three alternative formats of displaying information.

In Experiments 1 and 2, twenty-two certified flight instructors flew a general aviation simulator, either actively selecting and flying maneuvers to avoid traffic conflict by using a cockpit display of traffic information (CDTI) in a free flight experiment, or more passively executing corresponding maneuvers as commanded by a simulated air traffic controller, in a baseline experiment. The maneuvers spontaneously chosen by the pilots in the freeflight experiment were found to be only partially consistent with the maneuvers for traffic avoidance, dictated by current FAA "rules of the road", with pilots showing a marked preference for making vertical rather than lateral maneuvers, and climbing rather than descending maneuvers. Pilots also preferred to make simpler single axis maneuvers. Visual scanning in both experiments was measured, revealing that pilots spent a majority of their time fixating on the instrument panel (60% in baseline, 65% in freeflight). Furthermore, in free flight, they distributed their remaining visual attention differentially across the two sources of traffic information, the outside world and the CDTI. As a consequence, sometimes outside scanning dropped as low as 20%.

Traffic that was unknown to the data base generating the CDTI (i.e., in "mixed fleet" operations") did not appear to be hampered in its visual detection, relative to CDTI traffic, and relative to equally visible traffic in the baseline experiment, hence suggesting that the system is somewhat robust to imperfections in its data base.

In Experiment 3, 18 general aviation pilots flew a series of flight legs, similar to those of Experiment 1, but without traffic avoidance maneuvers. Three simulated datalink interfaces were used to convey instructions of varying message length, while pilots monitored for visual traffic. A synthesized voice display simulated the attention demands of the current baseline ATC system. A visual text display of instructions simulated many proposed data link formats. A redundant text-voice display was designed to address the memory demand limits of the auditory display and the attentional-head down limits of the visual display. As expected, communications (measured by readback errors) suffered with the auditory-only display, and also with greater message length. However, detection of traffic was **not** hindered by the visual text display, as pilots benefited by not having to write clearances down on a clipboard, an action they needed to take in the auditory-only display condition. Flight path tracking was best supported by the visual only data link display. Surprisingly, the redundant condition did not yield better performance on any task (communications, traffic scanning, flight path tracking), than either of the single

modality conditions. The results are interpreted within the context of theories of attention, using the visual scanning measures to help identify the influence of different attention components.

INTRODUCTION

Two new technologies appear to be on the verge of introduction into the 21st century cockpit: Data Link (Kerns, 1999, Navarro & Sikorski, 1999) is designed to improve the reliability of air-ground communications, while the Cockpit Display of Traffic Information (CDTI) is designed to support better traffic awareness for the pilot, as well as, potentially, provide a key component of the system necessary to support freeflight (Wickens, Mavor, Parasuraman, & McGee, 1998). Both technologies have implications for pilot attention and information processing capabilities. These implications may be amplified in general aviation, sometimes involving a single pilot, relative to their status in commercial aviation, where there are normally at least two pilots available to share the processing demands. In the following report, we first describe our research on the CDTI and freeflight, and then describe the research regarding data link. A common element in both domains is our interest in the allocation of visual attention to the outside world for traffic monitoring, a task that can be deemed of particular importance in general aviation AND in a freeflight environment (where ATC does not provide a redundant backup). Hence in both domains we measure visual scanning, and in both, we examine the effectiveness of such scanning, in finding traffic in the visual world.

FREEFLIGHT AND THE CDTI: EXPERIMENTS 1 AND 2

The new technology of the cockpit display of traffic information (CDTI), made possible by advanced communications, navigation and surveillance equipment, when coupled with the pilots' possible increased responsibilities for maneuvering to avoid conflict under the procedures of freeflight, imposes a host of potential changes to the pilots' tasks and cognitive/information processing operations. In the first two experiments, we focus on three related issues concerning these changes. First, we address the nature of the inherent choices or **decisions** that pilots make to maneuver to avoid traffic, when such maneuvers are supported exclusively by a CDTI (i.e., air traffic control is not involved). Second, we address the visual attention demands of pilots when engaged in such maneuvering. How much do they attend to the CDTI, to the instrument panel, or to the outside world? Finally, we examine the consequences of unreliable CDTI information that might result in an airspace in which not all aircraft have necessary equipage to communicate their position.

The first issue, the identity of pilots' inherent tendencies to chose one sort of maneuver over another to avoid conflict, is of critical importance as new "rules of the road" for free flight may be considered. It would seem essential that such rules conform to, or at least do not oppose, pilots' natural tendencies because otherwise the rules, whether followed from memory, or embodied in software decision aids that recommend certain maneuvers, might on occasion, be contradicted, leading to confusion and ambiguity in resolving a potential conflict (Pritchett & Hansman, 1997; Ciemer, Gambarani, Sheridan, Stangle, Stead, & Tillotson, 1993). While the FAA, through FAR #91.113 dictates a small number of existing rules of the road for conflict avoidance (e.g., lateral maneuvers are recommended, and these advise turning right), only a handful of studies have actually examined how pilots chose to maneuver to avoid a conflict, and surprisingly, these suggest some inherent conflict avoidance stereotypes that may violate the

FAR. For example, Beringer (1978) found that pilots tended to turn left, nearly as often as turn right. Merwin, Wickens, and O'Brien (1998) found an inherent bias to maneuver vertically, rather than laterally (as the latter is recommended by the FAR), a tendency that may reflect the greater simplicity and time efficiency of vertical over lateral maneuvering (Krozel & Peters, 1997; Helleberg & Wickens, 1999). Other researchers have found that the maneuver preferences are modulated, both by the particular conflict geometry (pilots tend to climb more when the traffic is descending from above, than when ascending from below; Merwin et al., 1998; and they tend to turn toward the traffic), and by the particular format of the display (vertical maneuvering is encouraged, to the extent that the display depicts altitude, either in perspective; Ellis, McGreevy, & Hitchcock, 1987; or particularly as a plan view; Merwin et al., 1998).

A shortcoming of these studies is that none have been carried out in the full cockpit environment typical of much of general aviation flight; that is, with the complete configuration of instrument panel, forward view and CDTI, the configuration that is used in the present experiment.

The second issue we address concerns the attentional requirements of the new responsibilities associated with self-separation. On the one hand, the cognitive demands of deciding how to maneuver, may interfere with (and disrupt) other cognitive aspects of flying, such as following procedures or monitoring. On the other hand, the visual attention demands of monitoring and processing information on the CDTI will clearly disrupt scanning of either the instrument panel, the forward view, or both. Although it may appear superficially that monitoring the outside world (for traffic) can be sacrificed given that such traffic now appears on the CDTI, such an assumption is only valid to the extent that all nearby traffic is represented on the CDTI. However, since current CDTI design assumes active transponders on such nearby aircraft, the CDTI will not be able to represent aircraft that are not so equipped. These can be seen only through the forward view; and unlike conventional IFR procedures, where these would be noted by air traffic control, in a freeflight environment, the pilot might be solely responsible for their detection, and this could be accomplished only by forward viewing.

Some previous research has examined the workload (attentional) implications of CDTI based free flight (Kreifeldt, 1980), and has inferred visual attention demands (Morphew & Wickens, 1998), while measures of pilot visual workload (head down time) associated with other non-CDTI technology have also been obtained in other research (Wreggit & Marsh, 1998). However no study has directly examined visual scanning imposed by use of the CDTI to maintain self separation. Finally, although some studies have examined pilot response to unexpected events when automation fails (Beringer & Harris, 1999), observing such responses to be disconcertingly long, no research has done so when the event to be detected is the appearance of unannounced traffic.

Thus the purpose of the first two experiments that we report here is fourfold:

1. To understand the basic pilot maneuver "stereotypes" or tendencies to maneuver to avoid conflict in a freeflight scenario.

2. To model the visual attention demands across three areas of interest: the CDTI, the instrument panel, and the outside world.

3

3. To evaluate how these demands are **changed**, by changing procedures from the current ATC-dominated airspace, to a free flight environment.

4. To establish the pilots' ability to notice aircraft, whose existence is not known by the CDTI equipage.

In order to address these goals, two experiments were conducted whose data will be described below. In a "baseline" experiment, pilots flew a series of legs in which traffic avoidance maneuvers were commanded by simulated ATC. In a free flight experiment, similar maneuvers were flown, but self selected by pilots from dynamic traffic information provided on the CDTI. In order to insure that maneuver type and complexity (and therefore instrument panel demand) was roughly equivalent between the two experiments, the baseline experiment was conducted **after** the freeflight experiment, and maneuvers for the baseline were constructed to mimic those demonstrated by freeflight pilots in their self-selections. Greater details regarding both experiments can be found in Wickens, Helleberg, and Xu (1999), Helleberg, Wickens, and Xu (2000) and Wickens, Xu, Helleberg, Carbonari, and Marsh (2000).

Freeflight Experiment: Methods

Twelve certified flight instructors flew a sequence of flight scenarios, over a period of three 2 hour sessions. Each scenario was flown in the left seat of a Frasca single engine light aircraft simulator, with a CDTI mounted to the left of the control yoke. A forward view graphics could depict out -of-cockpit viewing (including visual traffic within 5 miles), across a visual angle of 135 degrees (Figure 1). Each scenario consisted of ten 5-7 minute legs, whose required altitude, heading and airspeed was annunciated at the beginning of each leg (Figure 2). During 6 of the legs (conflict trials), a traffic intruder would appear, along with a second traffic aircraft that did not present a conflict (e.g., flying parallel, diverging, or at a different altitude). On the remaining four legs, only a nonconflict aircraft was present. On one of these four nonconflict trials, the traffic was a "transponder-off" aircraft (i.e., the aircraft not represented on the CDTI). For all aircraft, pilots were requested to call out "traffic in sight" when the traffic became visible in the outside world. The time for this callout was recorded. Conflict traffic could appear randomly from the right, left, above, below, behind (overtaking), crossing, or in front (converging). Pilots were instructed to maneuver, if necessary, in such a way as to avoid a loss of separation from any traffic (1.5 miles or 1000 vertical feet), but otherwise to minimize the overall departure from the flight parameters designated at the beginning of the leg. The experimenters did not stress the desirability of any particular maneuver.

Freeflight Experiment: Results

Across all pilots and legs, there were a total of 432 valid conflict legs. Based upon analysis of the continuously recorded flight and control parameters, we classified the maneuver types chosen spontaneously by the pilots. In addition, each maneuver and associated type, could be scored in terms of its safety, as operationally defined here by the time that the aircraft spent in a state of "predicted conflict" (a state in which a loss of separation would occur within 45 seconds if no further maneuver were undertaken). This state was indicated by a color change on the CDTI symbology, and pilots were explicitly instructed to avoid its occurrence. We highlight below the most important findings, with details available in Helleberg et al. (2000).

4

Figure 1.

Typical ATC command

"Cessna 1851 Zulu, Fly Heading 070, Climb and Maintain 5500, and 140 Knots."

Typical Cross-country Scenario

Includes:
6 conflict legs
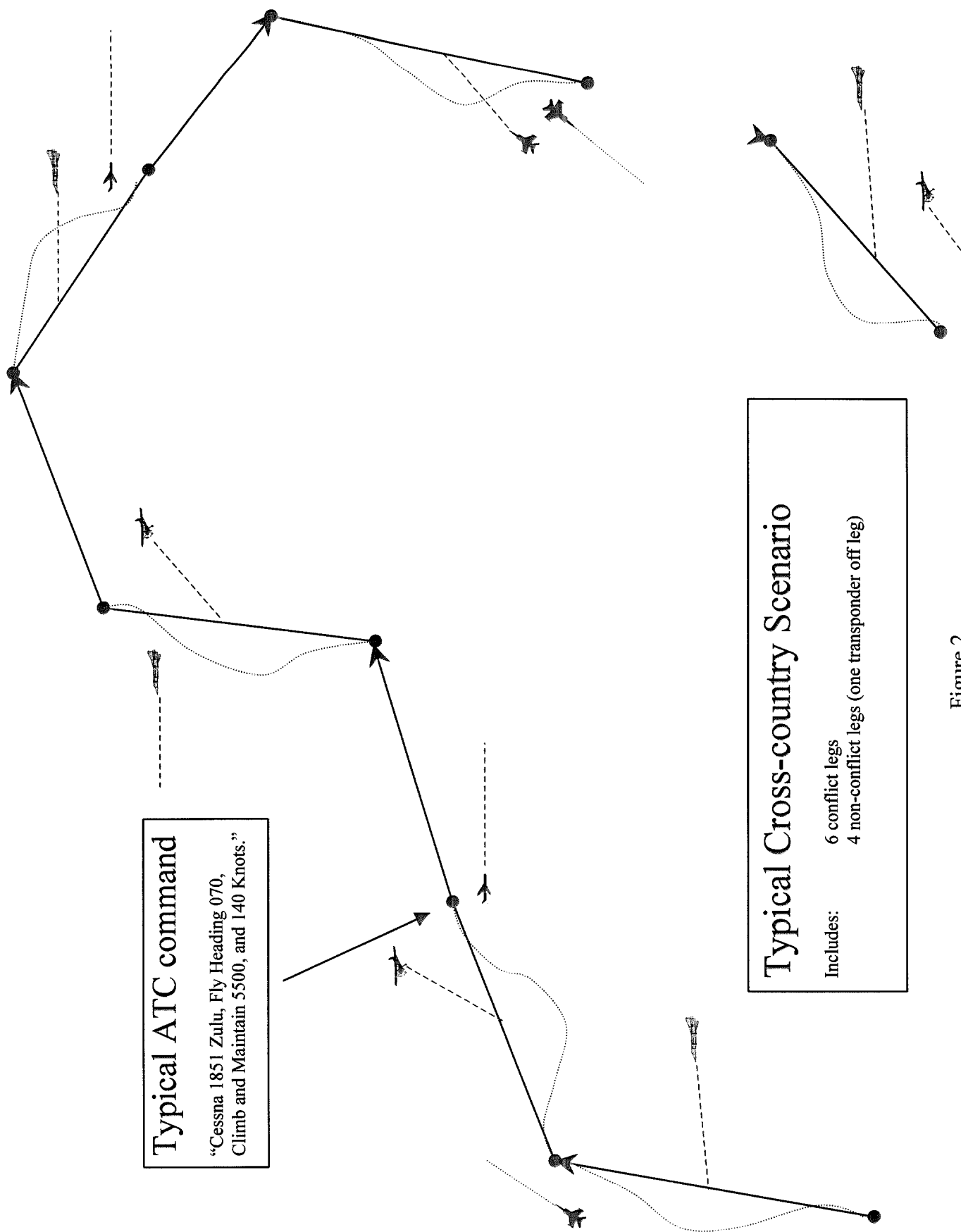4 non-conflict legs (one transponder off leg)

Figure 2.

1. Single axis maneuvers (92%) were vastly preferred over multi-axis maneuvers (e.g., combined airspeed and lateral change; 8%). Chi sq=92.87; p<.01.

2. Vertical maneuvers (75%) were far more frequently selected than lateral maneuvers (22%) and particularly than airspeed maneuvers (3%). Chi sq=85.62; p<.01.

3. When pilots chose to maneuver vertically, they were nearly twice as likely to climb as to descend. Chi sq=26.84; p<.01.

4. When pilots chose lateral maneuvers, there was no preference to turn right or left.

In addition to the above general preferences, the specific behavior of the traffic, visible on the CDTI also influenced the choice as follows:

5. Pilots maneuvered vertically in a direction opposite the vertical trend of the traffic (e.g., if the traffic was descending from above, pilots tended to climb, not descend). Chi sq=43.75; p<.01.

6. When maneuvering laterally, pilots tended to turn way from the traffic (57%) more than toward the traffic (43%). X^2=2.96; p=.09.

7. Pilots turned left (63%) more than right (37%) when traffic was converging head on, but reversed this tendency when traffic was crossing or overtaking (and therefore less visible in the outside world). Chi sq=5.39; p=.07.

We also examined the extent to which favored maneuvers were also **safer** maneuvers, as defined by the time avoided in a state of predicted conflict. This association of preference with safety was only partially confirmed. Thus while the favored single axis maneuvers were safer than the less favored multiaxis maneuvers (1. above F=3.14; p=.08), and while the favored climbing maneuvers were found to be significantly safer than the less favored descending maneuvers (3. above; t(291)=2.01; p<.01) the more favored vertical maneuvers were significantly less safe than the less-favored lateral maneuvers (2. above, t(355)=4.28, p<.01).

Finally, we analyzed the time required for the pilot to call out "traffic in sight". Most importantly, on non-conflict legs, there was an 18.8 second delay in the call out, if the traffic was **not** present on the CDTI (i.e., the "transponder off" aircraft; m=38.9 sec) relative to the CDTI-rendered aircraft (m=20.1 sec). However, the transponder-off aircraft were visually detected more accurately (69%), than those which were represented on the CDTI (52%).

Baseline Experiment and Scanning Analysis

As noted, maneuver behavior recorded in the freeflight experiment was used to construct a set of traffic avoidance instructions for traffic maneuvers, to be given by a simulated air traffic controller in the baseline experiment, such that the relative frequency of different maneuver types described above, was preserved. The ten pilots who participated in the baseline experiment followed the same general schedule and procedures as those in the Freeflight experiment. They did not have access to a CDTI, but could spot traffic visually in the same manner as the free flight pilots. Visual scanning was recorded for pilots in both groups. Because of problems in data

collection, the data reported below are based only on seven pilots from the free flight group. For the baseline pilots, our interest centered on the percentage of time spent, and mean dwell duration within, the two areas of interest (AOI) defined by the outside world (OW) and the instrument panel (IP) (We did not differentiate different instruments within the IP). For the Freeflight pilots, the CDTI was added as a third area of interest. More details of this analysis can be found in Helleberg, Wickens, and Xu (2000).

Table 1a shows the distribution of the percent time that the eye spent within each of the areas of interest for the baseline experiment (top row) and the freeflight experiment (bottom row), for all of the traffic conflict trials, in which an avoidance maneuver was necessary. Table 1b shows the same data for the non-conflict trials.

Table 1. Percent dwell time in each area of interest.

Experiment	IP	CDTI	OW	IP	CDTI	OW
Baseline	63	XXX	37	60	XXX	40
Freeflight	55	25	20	60	14	26

(a) conflict trials	(b) non-conflict trials

Statistical analyses, confirming the pattern of scanning in the table reveals that, for both experiments, whether in conflict or not, the instrument panel is the dominant focus of visual attention, demanding visual attention 61% of the time in the baseline experiment, and 56% of the time in the freeflight experiment. In contrasting the two experiments between the top and bottom rows of the table, we see that when self separation responsibility is imposed in the free flight experiment, on the conflict trials, pilots "borrow time" from both the IP and the OW to look at the CDTI, but appear to borrow more time from the OW (17%=37-20) than from the IP (8%=63-55). As a consequence in freeflight, the outside world is only scanned 20% of the time. (In fact, subsequent analysis reveals that this time drops to as low as 14% during that portion of a conflict trial when the pilot is the CDTI to plan a maneuver.) On non-conflict trials (Table 1b), pilots borrow time **exclusively** from the OW to look at the CDTI, as if they are trading off these two sources of traffic information, while preserving a full 60% of their viewing time to scan the instrument panel.

The mean dwell durations on the three AOIs are shown in Table 2, using the same format as Table 1.

Table 2. Mean dwell time (sec).

Experiment	IP	CDTI	OW		IP	CDTI	OW
Baseline	6.9	XXX	2.7		6.1	XXX	3.1
Freeflight	4.0	2.0	1.7		4.9	1.5	2.0

        (a) conflict trials   (b) non-conflict trials

While dwells are consistently longest on the instrument panel, which is the source of greatest "information" in the formal sense of the term (Senders, 1964), a comparison of the top and bottom rows of the two tables reveals that the mean dwell time on both the IP and the OW are shortened substantially (and significantly) by the free-flight-imposed requirement to attend to the CDTI. Thus, with freeflight, pilots engage in more rapid scanning between all AOIs, dwelling for shorter durations on all of them. Finally, both tables together suggest that on conflict trials the CDTI is the dominant source of **traffic** information, (both longer dwells and a greater percentage of time), whereas on non-conflict trials, the outside world is the dominant source.

We also examined the transition matrix between successive fixations in the free flight data to reveal where the scan would travel, after leaving a particular area of interest. Not surprisingly, on both conflict and non conflict trials, when the scan left either the CDTI or the outside world, it traveled most frequently to the dominant "attention sink" of the instrument panel. However when the scan left the IP, on conflict trials it would travel more frequently to the CDTI (t = 8.00, p<.01), but on non-conflict trials, it would travel more frequently to the OW (t=4.88, p<.01).

The more frequent cross scanning between the CDTI and the IP on conflict trials revealed by the transition analysis, would suggest that there might be a longer time that the outside world was left **unattended** on such trials, if the dwell time on the CDTI and IP was not shortened to compensate. This measure, referred to as the **mean first passage time** (MFPT; Moray, 1986) was calculated for the OW (see Wickens et al., 2000), and revealed the MFPT to the outside world was slightly longer on conflict trials (5.52 sec) than on non-conflict trials (4.41 sec). However, these two MFPTs were significantly less in free flight than their value in the corresponding conditions of the baseline experiment (6.9 sec and 6.1 sec) suggesting that pilots engaged in a relatively optimal adjustment of scanning: knowing that ATC was not responsible for far domain traffic monitoring in the freeflight experiment, they left the far domain unattended for shorter durations, even as their overall visual attention demands downward were increased by the CDTI.

It was important to establish whether the reduced amount of head up time, imposed by the CDTI, rendered the free flight pilots less effective in detecting and calling out visual sightings of the traffic. Such traffic was classified into three categories. (1) On conflict trials, this was the conflict aircraft (since there was no other traffic made visible). For the freeflight pilots, on non-conflict trials, most traffic, eventually visible in the outside world, was (2) initially visible on the CDTI because in the scenario, its transponder was assumed to be "on". (3) One

traffic aircraft out of 10 had its transponder "off" and hence was visible only in the outside world. For the baseline pilots, the distinction between "transponder on" vs. "transponder off" was not meaningful to the pilots, since there was of course no CDTI, However this distinction was important for data analysis, because the small subset of "transponder off" aircraft in Freeflight, had the same visibility or salience characteristics in the baseline experiment, thereby influencing their detectability. As it turned out, through purely random assignment, this small subset of transponder off trials was less salient, and less visible (leading to longer RTs) than the larger set which appeared on the CDTI.

Table 3 presents that callout times for the three classes of trials in the two experiments, showing in parenthesis the standard error and the percentage detected. The data show that conflict aircraft were detected visually more rapidly in the baseline than in the freeflight condition. This is neither surprising nor disturbing, since pilots were very much aware of the aircraft from the CDTI. The 4 second cost to detection in freeflight probably represents the added demands of processing the CDTI, and using it to compute the maneuver. More importantly, in the middle column (non-conflict trials), freeflight pilots were MORE effective than baseline pilots in detecting OW traffic, in spite of the fact that, as shown in Table 1, they were allocating less attention outside. Presumably they were doing so in a more vigilant or effective fashion, given their heightened responsibility for traffic monitoring (i.e., free-flight means no ATC). Furthermore, they were presumably assisted because the CDTI provided some guidance of where to look to visually sight the traffic. Most critically the data address whether the pilots' reliance upon the CDTI regarding where the traffic might be, could lead to a failure to detect traffic that was not on the CDTI (transponder off). These data, shown in the right column suggest no such failure. In freeflight these aircraft were detected 10 seconds faster than the equivalently positioned aircraft in the baseline experiment (36 vs. 46 seconds), again presumably reflecting the heightened vigilance and more effective OW scanning shown by the free flight pilots. While this advantage was not statistically significant, the finding allows us to reject the possibility that pilots were complacent in relying solely upon the CDTI to support their traffic awareness.

Table 3. Time to call out "traffic in sight" [standard error, accuracy]

	Conflict Trials	Nonconflict Trials	
Free Flight (N=13)	26 sec (1.01, 74%)	Transponder On (CDTI) 19 sec (2.13, 58%)	Transponder Off (OW only) 36 sec (4.20, 69%)
Baseline (N=10)	22 sec (1.03, 77%)	"Transponder On" 24 sec (2.82, 63%)	"Transponder Off" 46 sec (8.97, 67%)

Finally, it is important to note that the nature of the traffic avoidance maneuvers selected, discussed in the first part of the paper, had negligible influence on the scanning strategies that were deployed.

DISCUSSION: EXPERIMENTS 1 AND 2

The results and analyses of Experiments 1 and 2 reported here have revealed a complex, but readily interpretable picture of some of the consequences to pilot performance and cognition of introducing the responsibility for self separation, and the requisite monitoring of a CDTI. First, pilots appear to chose to maneuver in ways that only partially are consistent with existing FAR "rules of the road", and in some cases their preferences appear to contraindicate these rules. In particular they were seen to greatly prefer vertical maneuvering over the lateral maneuvering specified in the FAR, a preference based perhaps on the greater efficiency (Krozel & Peters, 1997) of that maneuver, but not its greater safety.

Second, in selecting the necessary information sources to engage in self separation, pilots appear to preserve a relatively optimal scanning strategy, conserving or exchanging visual attention between the two sources of traffic information (the outside world and the CDTI), as traffic monitoring and processing demands changed (conflict vs. non-conflict trials), but also preserving adequate monitoring of the outside world by making more frequent visits there, even as dwell times were shortened.

Third, the results suggested that scanning patterns adapted in free flight were able to preserve and in fact heighten the vigilance for detecting outside traffic, even when this traffic was not represented in the data base that generated the CDTI.

DATA LINK TECHNOLOGY: EXPERIMENT 3

The conventional proposal to provide datalink information in the cockpit has been to offer a text-based display in a head down location (Kerns, 1999, Navarro & Sikorski, 1999). Such a display will offer the "permanence" of information, in a way that buffers the vulnerabilities of working memory. At the same time, changes from an auditory to a visual display has several other implications for how pilot information is processed. For example, visual errors (confusing similar symbols like 5 and S), may replace auditory confusion, the visual data link may inhibit the sharing of party line information between aircraft (Midkiff & Hansman, 1992), or between pilots within an aircraft. Of particular interest in our studies are two somewhat competing attentional issues: visual resource competition, and auditory interruption.

On the one hand, it is clear that reading a digital datalink display **must** bring the pilot head down into the cockpit, in a way that can be avoided if the pilot merely listens to a voice transmission (whether from the air traffic controller over the radio, or from a voice synthesizer of a datalink message in the cockpit). This potential cost of head down time is consistent with an overall view of multiple resources (Sarno & Wickens, 1995, Wickens & Hollands, 2000), in which multiple task performance will be improved to the extent that tasks can be distributed between modalities, (auditory datalink with visual guidance and traffic monitoring) rather than focused exclusively within a modality (visual datalink, with visual guidance and traffic monitoring).

On the other hand, there is another factor related to "auditory preemption" that might mitigate the advantages of distributed resources (Latorella, 1998, Wickens & Liu, 1988). This is the observation that discrete auditory information, when it does arrive in the stream of an

ongoing visual task (such as traffic monitoring or instrument processing), will be more likely than discrete visual information to preempt the visual task. Thus, while the discrete information would be better processed if delivered auditorally than visually, the ongoing continuous visual task would be more disrupted by the discrete auditory than the discrete visual task. In short, auditory datalink could actually harm, rather than help, continuous visual monitoring, at least in the initial period following the arrival of a datalink message when this preemption would take place. Two explanations may be offered for this auditory interruption effect. First, the auditory modality is intrinsically more "attention grabbing", hence accounting for its preferred use as an alerting tool (Stanton, 1994). Second, because the auditory modality is transient, pilots may feel a need to deal with it immediately, before it is lost from working memory. The visual message, in contrast, can be dealt with more flexibly (Latorella, 1998), and its processing can be deferred until an ongoing visual task reaches a more "interruptible" period.

We also note a final factor that may in fact, offset any advantage of the auditory over the visual datalink, again regarding multiple resources. In a conventional auditory system, good pilots will adopt strategies to attempt to cope with their limited working memory. One such strategy is to write down key aspects of a message, a technique that, like the appearance of the text-based visual display will also bring the pilot's head into the cockpit, cutting into the time available for outside scanning.

It may be the case however that a **redundant** combination of both display modalities can capture the best of both worlds, in the same manner that redundant display has assisted in other domains (Wickens & Hollands, 2000, chapter 6). With such a combination, pilots can hear the message while looking outward, but look downward to check a message on the visual display if they believe that their working memory is faulty.

In the present experiment (see Helleberg & Wickens, 2000 for details), pilots flew a series of flight legs, responding to instructions of various length, and conveyed over either a visual-only, auditory-only, or redundant display. Pilots repeated the instructions, and readback accuracy was employed to assess communications performance. Traffic would periodically be visible in the outside world, forcing pilots to monitor outside in a way that might confer an advantage on the auditory display conditions. In the auditory-only condition, pilots were allowed to use a clipboard to take notes, in a way that is typical of their normal flight. However a separate condition was examined at the end of the experiment in which use of the clipboard was prevented. This was done in order to obtain a "pure" measure of the vulnerability of working memory, to the progressively longer messages.

METHODS: EXPERIMENT 3

Eighteen pilots, certified flight instructors from the same general population used in Experiments 1 and 2, volunteered to participate in Experiment 3. Each pilot flew in the simulator configuration shown in Figure 1, and flew a total of seven flight scenarios, similar to the example shown in Figure 2. Each scenario took approximately 25 minutes. Within each scenario, ATC instructions prior to each leg were delivered in one of three formats. An auditory format employed synthesized voice; a visual format employed a datalink display configured to the left of the instrument panel (the same location as the CDTI in the first two experiments), and a redundant format, presented both modalities. The ATC information presented at the beginning of

each leg varied systematically in its length, defined operationally in terms of the number of key elements that needed to be accurately perceived. These elements included heading, altitude, airspeed, a transponder squawk code, a barometer altimeter setting, and a radio frequency setting. Across all legs, the number of elements could vary from 2 to 6. Elements that did not change from a prior leg were not displayed. Pilots were instructed to read back the elements accurately as is their normal custom. In the auditory condition they were encouraged to use a clipboard as necessary. They were informed on-line of any mistakes in their readback, so that they would not implement an incorrect change to the flight path. As soon as they had read back the clearance, they were instructed to begin the maneuver (if any) to the new trajectory.

Traffic was potentially visible out the window at one of three times: immediately after the issuance of the datalink information, just after the information was estimated to have been processed (e.g., while pilots were beginning to implement any trajectory changes) and "well after" it was assumed to have been processed (e.g., while pilots were on a steady course). As in Experiments 1 and 2, traffic could appear at random locations on the outside screens, and required a "traffic in sight" callout. However this traffic was never designed to produce a conflict situation.

Following introduction and practice flights, the first six scenarios were used for data collection, and the order of formats was counterbalanced across pilots. Of these six, the last three employed eye movement measurement as in Experiments 1 and 2. Finally, a 7th scenario was presented with the auditory display, in which pilots were requested to rely **ONLY** on their memory for the readback.

RESULTS

Three independent variables were important in interpreting the results of the experiment: the display type, the length of the ATC message, and, for the traffic monitoring task, the **time** at which the traffic aircraft appeared, relative to the ATC message, whether during, just after, or well after the message was inferred to be processed. We describe the results below in order of the importance of the pilots' three subtasks in the order: aviate, navigate and communicate. Then we describe the effects on scanning strategy

Aviation: Flight Path Control

The results yielded clear evidence of a disruption of flight path control by the ATC messages. Across all three display conditions, longer ATC messages led to greater deviations in altitude, heading and airspeed. Disruptions on all three axes were however least with the visual display, and larger (and generally equivalent) with the two displays using the auditory modality (auditory-only and redundant). Only in airspeed error was there a trend for the redundant display to disrupt tracking less than the auditory-only display.

Navigation: Traffic Detection

The detection of traffic was clearly and understandably disrupted by the ATC messages. Those traffic aircraft appearing while a message was being processed were detected more slowly than those which appeared just after (and, generally, while the pilot was initiating the avoidance maneuver), and still slower than when the instructions arrived well after. Although there were

not strong effects of display format on aircraft detection, a weak trend (of marginal statistical reliability), observable only when the traffic appeared simultaneously with the message, suggested shorter RTs with the visual and redundant formats, than with the auditory-only format.

Communications: ATC Readback Errors

The analysis revealed a clear, and generally monotonic trend for more errors to occur with longer messages, observable across all three display formats. There was also a significant effect of format on readback errors, with greatest accuracy observed in the visual format (97%), intermediate accuracy in the redundant format (95%) and lowest accuracy in the auditory-only format (90%). The two independent variables did not interact. It should also be noted that the accuracy (80%) in the final no-clipboard trial of the auditory condition was greatly reduced, relative to the standard auditory condition.

Attention Allocation Strategies: Visual Scanning Analysis

Two parameters were critical in understanding how the different display formats modulated the allocation of attention across the pilot's visual world, (and in turn, how this modulation could help account for the results on the three primary tasks): the percentage of time the eye stayed within each area of interest (AOI), and the mean dwell duration within each AOI. An example of this scanning is shown in Figure 3, as it was recorded in the auditory condition.

Across all three display formats, the percentage dwell time (PDT) was dominated by the Instrument panel (approximately 60%), followed by the outside world (20-25%) followed by the **source of ATC information**. (10-15%). We defined this source to be the data link display for the visual and redundant conditions. In the auditory display, this source was defined as the clipboard, since this is where pilots directed their gaze while taking notes (to retain information for the longer messages). These fixation distributions were equivalent between the visual and the redundant conditions. However the auditory condition appeared to pull more attention away from the outside world, and allocate it to the ATC source (clipboard: 15%) than did the two visual conditions, in directing attention allocation to the datalink display (10%).

For all three displays, increasing message length fostered progressively more attention to the ATC source (datalink or clipboard), at the expense of attention to the OW. In contrast, attention allocated to the instrument panel was apparently "buffered" from any effect of increased message length.

The analysis of mean dwell duration (MDD) indicated that the instrument panel yielded longest dwells (approximately 4 seconds), considerably greater than the 1 - 2 second dwells that characterized all of the other AOIs. Dwells on the clipboard ATC source, for the auditory format (2 seconds), were considerably longer than dwells on the datalink display (1.2 seconds) as the ATC source for the visual and redundant formats. Finally, there was a very faint shortening of MDD across all three formats as the ATC messages increased in length.
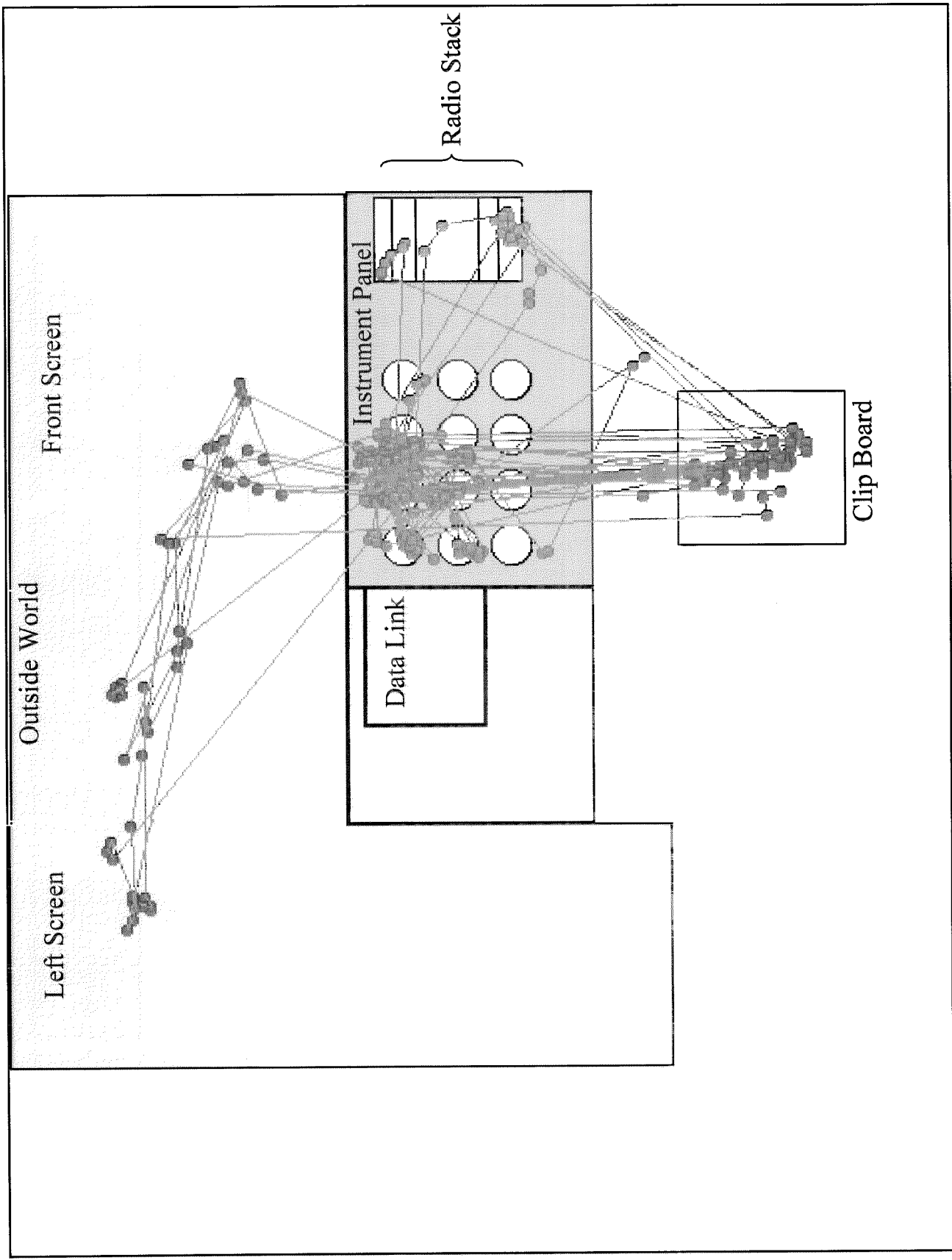
Figure 3.

DISCUSSION: EXPERIMENT 3

An overall comparison of the three key performance variables and the attention allocation visual scanning measure, as a function of the three display formats, is represented in Table 4, in which a "+" indicates that the display format supported good performance relative to the others, and a "–" indicates poor performance relative to the others. This judgment of superiority is based jointly on the statistical significance of the differences reported above, and upon their magnitude.

Table 4.

| | DISPLAY FORMAT | | |
	Aud	Red	Vis
AVIATION SUBTASK			
Communications	-	+	+
Attention allocation to ATC info.	-	+	+
Navigation (target detection)	-	+	+
Attention allocation to OW	-	+	+
Aviation (flight path control)	-	-	+
Attention allocation to IP	0	0	0
(0 indicates equivalent levels across all formats)			

In examining the table, three overall features are evident: first, the visual display format fairs best, the auditory format fairs most poorly, and the redundant format is intermediate. Second, for the first two aviation subtasks (communications and navigation), the visual scanning measures are consistent with the ordering of the display formats: greater allocation of visual resources leads to better performance. Third, this correspondence relation between scanning and performance is **not** observed for the Aviation subtask component. The overall pattern of effects revealed in the table can be interpreted within the context of three information processing mechanism as described below: working memory constraints, multiple resources and preemption or attention switching. We will use these mechanisms to account for performance on each of the three aviation subtasks.

Communications

The failure of the auditory display format to support good performance on the readback task, relative to the two formats using visual displays (redundant and visual) is not surprising, and indeed documents one of the major reasons why the datalink system has been advocated in the first place (Navarro & Sikorski, 1999; Kerns, 1999). That is, auditory working memory is vulnerable, and subject to forgetting. Less expected however was the failure of pilots to be able to use the clipboard and note taking facilities to fully compensate for the frailties of working memory with the auditory display. That is, despite this support, pilots still made more readback errors than when the visual display was present. Furthermore the use of the clipboard demanded more visual resources (2 seconds per head-down dwell) than did the readout of the visual datalink display (1.2 sec per dwell), a cost imposed on outside scanning to be addressed below.

One surprising finding was the failure of the redundant condition to offer the best of both modalities and hence provide superior performance to the visual format as has been observed in other domains (Wickens & Hollands, 2000). In fact, the level of accuracy was slightly lower in the redundant, relative to the visual format, a cost that may be attributed to pilots sometimes trusting their memory of the auditory instructions, without cross checking the visual display for assurance.

Navigation

We have defined the traffic callout as a navigation subtask because of the importance of hazard (traffic) awareness in safe navigation. It is apparent that this subtask suffered, as a consequence of the greater head down time and visual resource demand (relative to the datalink display) imposed by the clipboard note taking in the auditory-only condition. As noted above, dwells on the clipboard were longer than those on the visual display. Indeed the auditory display availed less scanning out the window (18%) relative to the two visual conditions (23%). The cost of this added head down time for the auditory format was only observed when it would be most expected; that is, when a target appeared just after a datalink message arrived, and therefore while that message was being processed. The RT cost for the auditory modality was not large, but was nevertheless observed in a direction that supports the visual display of datalink information, **relative to** conditions in which pilots must use note taking to compensate for auditory working memory constraints. As with the analysis of communications errors, we were somewhat surprised that the redundant format did not yield **better** performance (and more visual resources outside) than the visual format, given that the redundant format should allow pilots to continue their outside scan, and only look down to refresh their memory for the longer messages. It is possible that this failure of redundancy to offer advantages might have resulted from the **interruption** characteristics of the auditory display, described in the context of the aviate subtask as follows.

Aviation

We have operationally defined this highest priority task of **aviating** in terms of the pilots' ability to maintain the prescribed flight path in vertical, lateral and airspeed components, as achieved by allocating visual attention to the instrument panel. The results again suggest that the visual format supports best performance, but in contrast to the communications and navigation subtasks, the redundant format proved to be inferior, and led to performance more equivalent to the auditory format. Indeed such performance was fully equivalent to that with the auditory format in lateral and vertical control, and only in airspeed does it show a slight improvement. It is also evident that these differences in performance across all three formats are observed **despite** the essential equivalence of visual attention allocated to the instrument panel across the three formats (approximately 60%, a value very close to the values observed in Experiments 1 and 2).

We attribute this unique pattern of effects on the aviate subtask to an **auditory preemption** or interruption effect, by which discrete auditory tasks have a tendency to interrupt ongoing continuous visual ones. This effect had been described in basic dual task paradigms by Wickens and Liu (1988), and observed directly in a laboratory based datalink simulation by Latorella (1998). The fact that this pattern of disruption was **not** echoed by the visual attention allocation measure suggests that this preemption effect is a qualitatively different phenomenon

from the visual resource competition which underlay the disruption of visual target detection. Further evidence for the role of a distinct mechanism from visual resource competition is the fact that the preemption effect appears to be common to both formats that used the auditory display. The preemption effect may therefore represent more of a sudden **cognitive** switch, induced by the discrete appearance of the auditory message, (and imposing momentary deviation from the required flight parameters), than from a reallocation of **visual** attention away from the instrument panel.

In conclusion, it appears that the proposal for a visual datalink display is indeed one that supports best overall performance for general aviation pilots. Despite the requirement that this format imposes for head down activity, the permanence of the visual display allows pilots to allocate that head down attention diversion in a manner that is more flexible and less disruptive of ongoing visual tasks, than is required by the need to process (and take notes on) the auditory transmission of data. The failure of the redundant use of both channels to support "the best of both worlds" is surprising and would reflect either the large contribution of auditory perception, or the failure to rely sufficiently on the auditory channel when desirable. It is possible that such a format could provide advantages if training was offered as to the appropriate use of each modality, an issue that requires further investigation.

ACKNOWLEDGMENTS

REFERENCES

Beringer, D. B. (1978). Collision avoidance response stereotypes in pilots and nonpilots. Human Factors, 20, 529-536.

Beringer, D. B., & Harris, H. C., Jr. (1999). Automation in general aviation: Two studies of pilot responses to autopilot malfunctions. The International Journal of Aviation Psychology, 9(2), 155-174.

Ciemer, C. E., Gambarani, G. P., Sheridan, E. H., Stangle, S. C., Stead, R. P., & Tillotson, D. H. (1993). The results of the TCAS II transition program (TTP). The ARINC Research Corporation.

Ellis, S. R., McGreevy, M. W., & Hitchcock, R. J. (1987). Perspective traffic display format and airline pilot traffic avoidance. Human Factors, 29, 371-382.

Helleberg, J. R., & Wickens, C. D. (1999). Pilot Maneuver Choice and Safety in a Simulated Free Flight Scenario (John Helleberg's First Year Project at the University of Champaign at Urbana-Champaign).

Helleberg, J. & Wickens, C. D. (2000). <u>Datalink modality in general aviation</u> (ARL-00-7/FAA-00-4). Savoy, IL: University of Illinois, Aviation Research Lab.

Helleberg, J., Wickens, C. D., & Xu, X. (2000). <u>Pilot maneuver choice and safety in a simulated free flight scenario</u> (ARL-00-1/FAA-00-1). Savoy, IL: University of Illinois, Aviation Research Lab.

Kerns, K. (1999). Human factors in air traffic control/flight deck integration: Implications of data-link simulation research. In D. J. Garland, J. A. Wise, & V. D. Hopkin (Eds.), <u>Handbook of aviation human factors</u> (pp. 519-546). Mahwah, NJ: Lawrence Erlbaum.

Kreifeldt, J. G. (1980). Cockpit displayed traffic information and distributed management in air traffic control. <u>Human Factors, 22,</u> 671-691.

Krozel, J., & Peters, M. (1997). Conflict detection and resolution for free flight. <u>The Air Traffic Control Quarterly Journal Special Issue on Free Flight.</u>

Latorella, K. (1998). Effects of modality on interrupted flight deck performance: Implications for data link. <u>Proceedings of the 42nd Annual Meeting of the Human Factors and Ergonomics Society</u> (pp. 87-91). Santa Monica, CA: Human Factors and Ergonomics Society.

Merwin, D. H., Wickens, C. D.,& O'Brien, J. V. (1998). Display format-induced biases in air traffic avoidance behavior. <u>Proceedings of the World Aviation Congress</u> (98WAC-71). Warrendale, PA: Society of Automotive Engineers, Inc.

Midkiff, A. H., & Hansman, R. J., Jr. (1992). Identification of important "party line" information elements and the implications for situational awareness in the datalink environment. <u>SAE Technical Paper Series 922023</u>. Warrendale, PA: Society of Automotive Engineers, Inc.

Moray, N. (1986). Monitoring behavior and supervisory control. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), <u>Handbook of perception and performance, Vol II.</u> (pp. 40-1-40-51). New York: Wiley & Sons.

Morphew, E. M., & Wickens, C. D. (1998). Pilot performance and workload using traffic displays to support free flight. <u>Proceedings of the 42nd Annual Meeting of the Human Factors & Ergonomics Society</u>. Santa Monica, CA: Human Factors and Ergonomics Society.

Navarro, C., & Sikorski, S. (1999). Datalink communication in flight deck operations: A synthesis of recent studies. <u>International Journal of Aviation Psychology, 9</u>(4), 361-376.

Pritchett, A. R., & Hansman, R. J. (1997). Experimental studies of pilot performance at collision avoidance during closely spaced parallel approaches. <u>Proceedings of the 9th International Symposium on Aviation Psychology.</u> Columbus, OH: Dept. of Aviation.

Sarno, K. J., & Wickens, C. D. (1995). The role of multiple resources in predicting time-sharing efficiency: An evaluation of three workload models in a multiple task setting. <u>International Journal of Aviation Psychology, 5</u>(1), 107-130.

Senders, J (1964). The human operators as a monitor and controller of multidegree of freedom systems. IEEE Transactions on Human Factors in Electronics, HFE-5, 2-6.

Stanton, N. (1994). Human factors of alarm design. London: Taylor and Francis

Wickens, C. D., Helleberg, J., & Xu, X. (1999). Maneuver choice in free flight. Proceedings of the World Aviation Congress (1999-01-5591). Warrendale, PA: Society of Automotive Engineers, Inc

Wickens, C. D., & Hollands, J. (2000). Engineering psychology and human performance (3rd ed.). New York: Prentice Hall.

Wickens, C. D., & Liu, Y. (1988). Codes and modalities in multiple resources: A success and a qualification. Human Factors, 30, 599-616.

Wickens, C. D., Mavor, A. S., Parasuraman, R., & McGee, J. P. (Eds.). (1998). The future of air traffic control: Human operators and automation. Washington, DC: National Academy Press.

Wickens, C. D., Xu, X., Helleberg, J. R., Carbonari, R., & Marsh, R. (2000). The Allocation of Visual Attention for Aircraft Traffic Monitoring and Avoidance: Baseline Measures and Implications for Freeflight (ARL-00-2/FAA-00-2). Savoy, IL: University of Illinois, Aviation Research Lab.

Wreggit, S. S., & Marsh, D. K. II. (1998). Cockpit integration of GPS: Initial assessment—menu formats and procedures (DOT/FAA/AM-98/9). Washington, DC: Federal Administration Office of Aviation Medicine.